JKPK (JURNAL KIMIA DAN PENDIDIKAN KIMIA), Vol. 10, No. 2, 2025 — pp. xx-xx
Chemistry Education Study Program, Universitas Sebelas Maret — ISSN 2503-4146
https://jurnal.uns.ac.id/jkpk — ISSN 2503-4154 (online)

PROFILING MULTICOMPONENT CHEMICAL REASONING: A LEARNING ANALYTICS APPROACH TO APPLIED AND SOCIO-CHEMICAL DIMENSIONS

Nur Hidayat, Woro Sumarni*, Endah Fitriani Rahayu, Sri Kadarwati, Kasmui, Dimas Gilang Ramadhani

Chemistry Education, Faculty of Mathematics and Natural Sciences, Universitas Negeri Semarang, Semarang, Indonesia

ARTICLE INFO

Keywords:
Assessment profiling; Data visualisation; Learning analytics; Reasoning performance; Student clustering.

Article History:
Received: 2025-06-16
Accepted: 2025-08-15
Published: 2025-08-31
doi: 10.20961/jkpk.v10i2.10424

ABSTRACT

Scientific reasoning in chemistry involves the ability to apply conceptual knowledge in problem-solving, as well as to evaluate issues within broader social, ethical, and environmental contexts. However, conventional assessments often fail to capture this multidimensionality by reducing performance to a single final score. This study uses an integrated learning analytics approach to analyze students' reasoning performance across two core domains of chemistry learning—applied reasoning and socio-chemical reasoning. A quantitative descriptive design was employed, involving 56 pre-service chemistry teachers who completed four open-ended essay questions, two in each reasoning domain. Student responses were scored using an analytical rubric assessing conceptual accuracy, logical coherence, and justification relevance. Data were analyzed using single-domain and multicomponent strategies, including quadrant profiling, trajectory mapping, clustering, and distribution analysis. Visual tools such as radar charts, spaghetti plots, contour density plots, and alluvial diagrams were used to depict students' reasoning profiles. Results revealed that most students demonstrated moderate reasoning abilities, although notable inconsistencies were observed between the domains. Individual trajectories exhibited non-linear variations, highlighting diverse cognitive patterns. Clustering and heatmaps indicated distinct learner segments, while alluvial diagrams illustrated transitions between reasoning levels across domains. These findings suggest that students' reasoning abilities are varied and dynamic. It is concluded that chemistry reasoning is multidimensional and should be assessed through integrated, data-driven methods. The study recommends the adoption of formative, analytics-supported assessments to inform differentiated instruction and promote deeper conceptual and ethical engagement in chemistry education.

**Corresponding Author* woro@mail.unnes.ac.id
How to cite: N. Hidayat, W. Sumarni, E. F. Rahayu, S. Kadarwati, Kasmui, and D. G. Ramadhani, "Profiling Multicomponent Chemical Reasoning: A Learning Analytics Approach to Applied and Socio-Chemical Dimensions," *Jurnal Kimia dan Pendidikan Kimia (JKPK)*, vol. 10, no. 2, pp. xx–xx, 2025. [Online]. Available: https://doi.org/10.20961/jkpk.v10i2.104248 .

INTRODUCTION

Scientific reasoning is a fundamental component of science education, particularly in chemistry learning, as it plays a crucial role in fostering deep conceptual understanding and supporting evidence-based decision-making [1], [2]. In modern educational settings, students are expected not only to master chemical concepts symbolically but also to apply their knowledge meaningfully in real-world contexts that are socially and environmentally relevant [3]-[5].

Consequently, higher-order thinking skills (HOTS), such as conceptual reasoning and scientific justification, have become central to curriculum and assessment reforms both nationally and internationally [6], [7]. Despite this, many existing studies have examined reasoning in a fragmented and one-dimensional manner, neglecting that students often reason simultaneously across multiple domains by integrating macroscopic, microscopic, and symbolic representations in building chemical understanding [8], [9]. This highlights the need for a more holistic approach to chemistry assessment and instruction that evaluates outcomes and fosters analytical, reflective, and contextually grounded thinking through laboratory practices and problem-based learning experiences [10], [11].

In practice, students often demonstrate inconsistent performance in applied and sociochemical reasoning. Many students excel in procedural or symbolic tasks but struggle when required to engage with ethical, environmental, or societal dimensions of chemical issues [12], [13]. Standard assessments rely heavily on aggregated final scores, offering limited insight into students' performance patterns across different question types or reasoning domains [14]. As a result, overly generalised analyses fail to capture the heterogeneity of students' reasoning trajectories and the nuanced interconnections between cognitive domains [15], [16]. This challenges teachers who must design pedagogical interventions responsive to each student's cognitive profile. Accordingly, there is an urgent need for integrated and data-informed assessment models that map cross-domain reasoning performance and offer constructive formative feedback. Instruction that includes real-world contexts, social reflection, and multidimensional reasoning is more effective in supporting students' robust scientific reasoning skills development [17], [18].

Despite this need, a critical research gap in chemistry education exists in analyzing students' scientific reasoning performance, particularly in integrating multidimensional profiling. Most previous studies have focused on final scores as isolated achievement indicators, neglecting the interaction between cognitive domains such as applied and socio-emotional reasoning [19]. This limited focus fails to capture how students integrate knowledge across procedural, conceptual, and contextual dimensions, thereby reducing the diagnostic value of assessments. Moreover, learning analytics approaches such as quadrant visualisation, individual trajectory mapping, and clustering analysis are still rarely employed in chemistry education research, despite their potential to reveal detailed and personalised patterns of student reasoning. The absence of analytic techniques that map students' performance across score combinations (rather than averages) hinders a nuanced understanding of reasoning diversity within a cohort. Consequently, teachers struggle to design interventions that match the cognitive profiles of individual learners.

To address these analytical and pedagogical limitations, there is a growing recognition that data visualisation tools can play a central role in bridging the gap between traditional assessments and the complexity of students' reasoning. Recent advancements in science education research have increasingly applied data-driven methods such as cluster analysis, heatmaps, and multilevel modelling to investigate learning patterns at both individual and institutional levels [20]-[22]. This reflects a shift from static, summative assessments toward more diagnostic approaches that reflect the dynamic nature of scientific reasoning. Although multidimensional reasoning frameworks, including two-dimensional and multicomponent models, have gained theoretical traction, their practical application in chemistry education remains limited [23]. In contrast, fields such as mathematics and literacy have successfully implemented these frameworks to identify cognitive profiles and inform adaptive instruction [24], demonstrating their utility in enhancing learning outcomes. Chemistry education, however, continues to rely heavily on score aggregation, overlooking the diagnostic power of advanced analytics in revealing students' underlying cognitive structures [25], [26]. Thus, more comprehensive assessment models are needed to integrate learning analytics with reasoning frameworks to support evidence-informed, responsive pedagogy.

In response to these challenges and opportunities, this study introduces an integrated multidomain reasoning analysis that combines Applied Reasoning and Socio-chemical Reasoning through advanced data visualisation techniques. Unlike prior studies that relied on aggregated scores, this research conducts granular analysis at the item level and across individual reasoning trajectories, enabling the identification of nuanced performance patterns [27]. The study applies underutilised visual tools in chemistry education—including spaghetti plots, quadrant scatter plots, contour density plots, and alluvial diagrams—to uncover cognitive variability and cross-domain interactions in more interpretable forms [28], [29].

These tools offer methodological and pedagogical value by making complex learning data accessible and actionable. Spaghetti plots illustrate individual reasoning progressions over time, revealing dynamic fluctuations in student thinking [30], [31]. Quadrant scatter plots map performance across two reasoning dimensions, enabling the detection of reasoning asymmetries and student profiles that support differentiated feedback [32], [33]. Contour density plots reveal dominant response distributions and transitional zones, facilitating diagnostic insight into reasoning diversity [34]. At the same time, alluvial diagrams visualise shifts in reasoning categories, capturing how students transition across applied and contextual domains [35], [36]. Together, these tools support a data-informed approach to formative assessment and instruction. Furthermore, this study presents a practical model of how multicomponent profiling can inform pedagogical decision-making by identifying student needs and guiding targeted [37]. As a result, the research contributes methodologically and pedagogically to advancing visual, formative, and evidence-based assessment practices that promote scientific reasoning literacy in chemistry education.

Aligned with the identified gaps and emerging practices, this study aims to analyse students' performance across two critical domains of scientific reasoning: Applied Reasoning and Socio-chemical Reasoning. These domains capture students' ability to apply chemical concepts procedurally and engage in reflective, ethical reasoning on socially embedded issues [38]. By examining both, the study seeks to construct a nuanced understanding of students' cognitive functioning in chemistry, extending beyond symbolic mastery toward decision-making grounded in real-world contexts. To achieve this, the research adopts a multicomponent profiling approach to map combinations of student performance across domains, enabling the identification of reasoning patterns often obscured in traditional, unidimensional assessment models [39]. Furthermore, it investigates individual trajectories across tasks to uncover performance fluctuations and clustering tendencies, offering insight into student segmentation based on reasoning profiles [40]. Central to this endeavour is the use of advanced visual tools, which serve not only as analytic techniques but also as pedagogical instruments that support differentiated instruction and informed decision-making in classroom contexts [41]. Through the integration of domain analysis, trajectory mapping, and learning analytics, this study contributes to a more responsive and evidence-informed model of chemistry education that fosters holistic scientific reasoning.

METHODS

1. Research Design

This study used a quantitative descriptive approach and learning analytics methods to explore students' multicomponent reasoning ability profiles. The main focus of the research design was to analyse and identify patterns of student performance in two cognitive domains central to chemistry learning, namely Applied Reasoning and Socio-Chemical Reasoning. Applied Reasoning refers to students' ability to apply chemical concepts technically and conceptually in an academic context. At the same time, Socio-Chemical Reasoning encompasses the ability to reason about contextual chemical issues involving social, ethical, and environmental dimensions. These two domains were chosen because they reflected students' real challenges in integrating scientific knowledge into everyday situations.

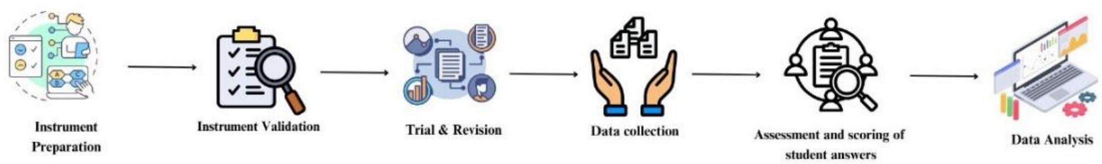

Figure 1. Research Design

This study employed a multi-layered analysis approach to gain a deeper understanding of cognitive variation between individuals and groups. This analysis included examining student performance at the individual level to explore their thought processes and the consistency of their responses across different questions. It also included collective-level analysis to identify common patterns in score distributions and domain combinations. Through this design, the study aimed to reveal students' final achievements and uncover the dynamics of their reasoning processes, which provided a foundation for developing more adaptive and responsive assessment and learning strategies.

2. Participant

The subjects in this study were 56 first-year pre-service teachers who had completed a series of reasoning-based tasks in the context of chemistry learning. Participants were selected purposively using strict inclusion criteria to ensure data validity. Only students who had completed all four reasoning questions in full, without any blank answers, and who had achieved full scores in the two main domains of Applied Reasoning and Socio-Chemical Reasoning were included in the analysis. These criteria were established to ensure that each participant provided representative data that could be analyzed comparatively across reasoning dimensions. All data used in this study were anonymized to protect the confidentiality of the participants' identities. Before the data collection, the researchers obtained official approval from the relevant institutions, per applicable research ethics procedures. This step ensured the entire research process was conducted ethically and responsibly, providing a safe participatory space for students to contribute to developing chemistry education studies.

3. Instrumentation

The main instrument used in this study consisted of four short essay questions developed using a tiered reasoning approach, or two-level reasoning. Each question evaluated the depth of students' thinking through conceptual understanding and reason-based justification. The first two questions (Questions 1 and 2) focused on measuring applied reasoning, which is defined as students' ability to understand and apply chemical concepts through symbolic representation and scientific logic. The next two questions (Questions 3 and 4) were designed to assess Socio-Chemical Reasoning, which included interpreting chemical problems in a social and ethical context and making decisions based on considerations of their impact on society and the environment.

Each question was scored using an analytical scoring rubric ranging from 0 to 4, based on three main indicators of reasoning: conceptual accuracy, logical coherence, and the justification and relevance of the reasons provided by the student. With two questions in each domain, the maximum score that could be obtained was 8 per reasoning domain, resulting in a total score of 16 points. This scoring structure allowed for a comprehensive and balanced measurement of student performance across both technical and socio-chemical contexts, while providing a solid quantitative basis for further analysis of reasoning profiles.

4. Data Collection and Scoring

The data collection process in this study was conducted through documentation of students' written responses to four reasoning questions designed around two core cognitive domains: Applied Reasoning and Socio-Chemical Reasoning. Each student completed the tasks individually in a short essay format, without strict time constraints, allowing for the elaboration of reasoning depth and clarity. Student responses were subsequently evaluated using a rubric-based quantitative scoring system developed with reference to established educational frameworks to ensure rigour, validity, and alignment with current practices in chemistry education research.

The rubric incorporated four key indicators: mechanistic and causal reasoning, quantitative model-based reasoning, contextual and representational application, and risk–benefit evaluation. Mechanistic and causal reasoning was informed by the work of Heisterkamp and Talanquer (2015), who emphasised that articulating causal mechanisms enhances students' conceptual understanding of chemical phenomena. Quantitative model-based reasoning drew upon Gilbert and Justi's (2016) framework on modelling in science education, reflecting students' ability to apply mathematical or symbolic representations in constructing scientific arguments. The contextual and representational application indicator was adapted from contextual learning approaches such as those discussed by Dewi and Primayana (2019), assessing how students related chemical ideas to real-life contexts using appropriate scientific representations. The risk–benefit evaluation indicator reflected the growing emphasis on socio-chemical reasoning. It was grounded in frameworks advocating the integration of ethical and environmental considerations into science education, as highlighted by Mendonça and Justi (2013). Together, these indicators provided a comprehensive and theoretically grounded framework to assess technical and socio-contextual dimensions of students' reasoning, thereby supporting a valid, reliable, and educationally meaningful evaluation of their chemical reasoning.

To maintain the objectivity and consistency of the assessment results, the scoring process was carried out by two independent assessors with backgrounds in chemistry education and experience in reasoning-based assessment. The validity of the scoring results was strengthened through interrater reliability testing, which in this study was calculated using Cohen's Kappa coefficient. This approach was chosen to ensure that the interpretation of student responses was not subjective and that the assessment of reasoning indicators was conducted consistently. The high reliability test results formed the basis for the validity of the evaluation process and provided confidence in the quality of the data used in further analysis.

5. Data Analysis Procedure

The instruments used in the research were (1) an online research platform using webcam eye-tracking (RealEye); (2) computer devices; (3) chemistry problem-solving questions consisting of 2 main questions with six sub-questions each, arranged according to the problem-solving steps; (4) interview guidelines.

6. Ethical Consideration

This study was conducted per the ethical principles of educational research to ensure the integrity of the process and the protection of participants' rights. Before data collection, the researcher obtained official approval from the relevant educational institutions, including the program administrators and the authorized academic ethics unit. All data collected from participants were anonymized, so that students' personal identities could not be identified either in the analysis process or in the reporting of results. This anonymization included removing names, student ID numbers, and other sensitive information from answer sheets and databases.

Furthermore, participation in this research had no implications for students' academic assessments. Students were informed openly about the purpose and scope of data use and were assured that their participation would not influence their academic outcomes. This process was conducted transparently before data collection, with researchers providing both verbal and written explanations that included the right to opt out or withdraw at any time without consequence. Through this approach, the study ensured compliance with academic ethical standards and created a safe, voluntary, and trust-based environment for student participation.

RESULT AND DISCUSSION

1. Distribution of Applied and Socio Scores

The initial analysis focused on the distribution of students' scores in the two domains of chemical reasoning, namely Applied Reasoning and Socio-Chemical Reasoning. As illustrated in Figure 2, the histogram on the left shows that Applied Reasoning scores exhibit a right-skewed distribution. Most students scored between 5 and 7, with a mean of 5.72 and standard deviation of 1.81, indicating a concentration in the moderate to high proficiency range. However, a group of students fell in the lower tail (scores 1–2), suggesting a substantial performance gap. In contrast, the histogram on the right representing Socio-Chemical Reasoning showed a more symmetrical distribution, with a concentration of scores in the 4–5 range and a smaller standard deviation (1.45), reflecting more uniform performance in this domain.

This divergence in score distributions suggests different cognitive and instructional influences underlying the two reasoning domains. The broader spread of applied reasoning may be due to its alignment with traditional teaching methods emphasizing algorithmic and procedural problem-solving using chemical equations and symbolic representations. These methods often advantage students with stronger backgrounds in mathematics and formal logic

[42]. In contrast, the tighter clustering of Socio scores implies a shared challenge among students in articulating value-based or ethically grounded reasoning, possibly due to limited exposure to socioscientific issues in standard curricula [43]. Thus, while the technical domain shows stratification, the contextual domain highlights a systemic instructional gap.

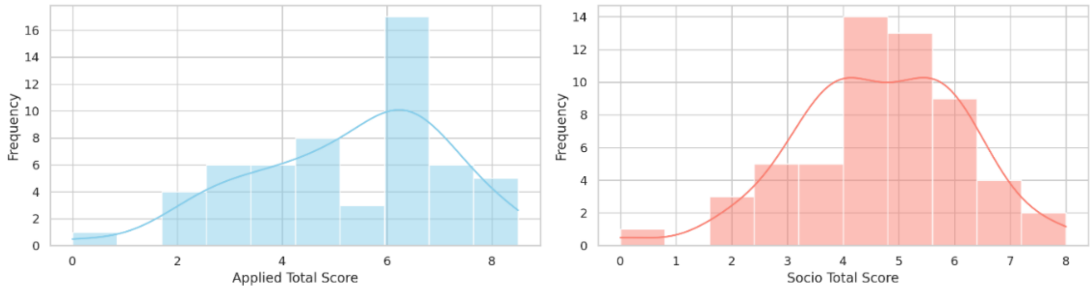

Figure 2. Histogram of Applied and Socio-Chemical Reasoning Scores.

From a pedagogical perspective, these findings underscore the need for a more integrated curriculum that addresses procedural fluency and contextual reasoning. Current educational reforms stress the importance of cultivating 21st-century skills such as ethical reasoning, systems thinking, and decision-making under uncertainty—skills encapsulated within Socio-Chemical Reasoning [44]. The fact that some students perform well in Applied Reasoning but underperform in Socio-Chemical Reasoning suggests that cognitive proficiency in one domain does not ensure competence in the other. Therefore, curriculum strategies such as project-based learning, scenario analysis, and value clarification must be systematically embedded in chemistry education to close this reasoning divide.

2. Quadrant Profile of Combined Reasoning

The analysis of student performance across the two reasoning domains, Applied Reasoning and Socio-chemical Reasoning, revealed important patterns of cognitive variability, as visualised in Figures 3 and 4. The contour plot in Figure 3 showed a high concentration of student scores around an Applied score of 6 and a Socio Chemical score of 5, indicating that most students performed moderately in both domains. This cluster corresponded to mid-level indicators in the rubric, where students demonstrated basic conceptual accuracy and procedural logic in the applied domain, yet showed limited integration of ethical or contextual reasoning in the socio-chemical domain. These responses often included partially correct chemical representations and acknowledgment of social issues but lacked deep causal justification or evidence-based evaluation. Regarding rubric alignment, the dense region of the plot reflected moderate-level performance primarily characterised by accuracy in procedural application and general contextual awareness, but with limited evidence of mechanistic reasoning, model-based support, or ethical evaluation. The low-score zones aligned with responses that lacked all four indicators: superficial explanations, lack of symbolic or model-based reasoning, and minimal socio-contextual engagement. Conversely, the sparsely populated high-score regions corresponded to responses demonstrating integrated reasoning across all indicators, including causal mechanisms, quantitative modelling, appropriate contextual representations, and risk-benefit justification.

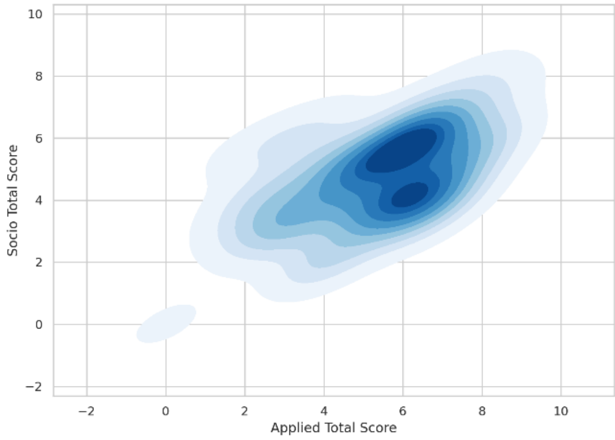

Figure 3. Contour Plot of Combined Reasoning Scores

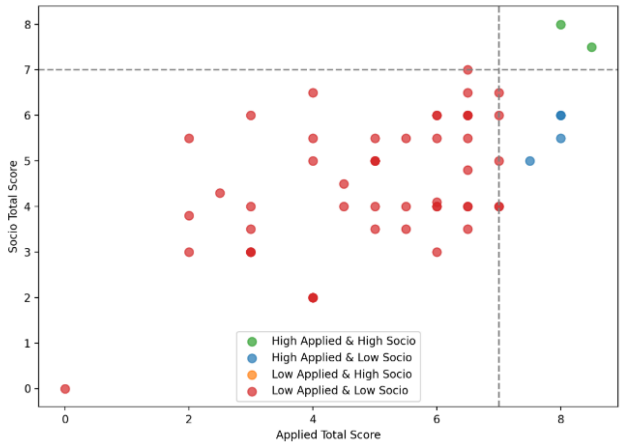

Figure 4. Quadrant Mapping of Multicomponent Reasoning Profiles

Lower-density regions of the plot reflected students who either underperformed or excelled across both domains. Students in the low-score region, with Applied and Socio scores below 4, tended to exhibit fragmented reasoning, minimal use of models, and superficial references to context. In contrast, the high-score region, with scores above 7 in both domains, represented students whose responses aligned with the highest-level rubric indicators such as strong causal explanations, model-supported reasoning, and thoughtful consideration of ethical implications. These findings suggest that high-level multicomponent reasoning remains rare, while moderate procedural reasoning is more prevalent.

The quadrant map in Figure 4 further illustrates the distribution of students based on their combined domain performance.

3. Individual Reasoning Trajectories

Only a small number of students, fewer than five percent, were categorised as High High, indicating balanced reasoning across all dimensions. Nearly twenty percent fell into the Low Low group, revealing widespread difficulty in both technical and contextual reasoning. Most students in the High Low quadrant showed strong procedural thinking but weak socio-contextual reasoning. This imbalance reinforces prior research that science education often prioritises technical mastery while overlooking ethical and contextual dimensions [6]. As observed by Dewi and Yahdi [45], this imbalance may result in students being prepared to solve textbook problems but underprepared to respond to real-world chemical issues.

The scarcity of High High performers highlights a gap in instructional models that do not adequately foster the development of scientific and ethical reasoning skills. Addressing this challenge requires chemistry education to go beyond content delivery and encourage students to reflect on the broader impact of scientific practice [43]. As [46] This development demands learning experiences connecting disciplinary knowledge with social and environmental contexts. Introducing authentic case studies such as climate change, pollution management, and decision making in chemical industries can support the formation of well-rounded reasoning and prepare students to apply chemistry in scientifically sound and socially responsible ways [47].

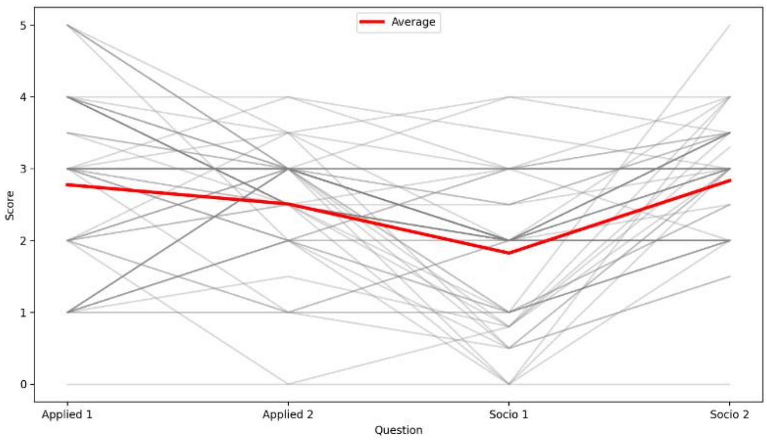

Figure 5. Spaghetti Line Plot of Individual Reasoning Trajectories.

Individual analysis of student performance trajectories visualized in Figure 5 (Spaghetti Line Plot) shows significant variation in how students respond to four reasoning questions covering two domains. The pattern shown generally indicates fluctuations in scores between questions, with a sharp decline from Applied Reasoning to Socio-Chemical Reasoning. Socio 1 is the question with the lowest average score, which may be influenced by the order in which the questions are presented or the increased cognitive demands due to a more complex social context [48]. The steeply declining trajectory lines reflect a cognitive gap between students' procedural and reflective abilities. This indicates that success in solving applied questions does not necessarily translate into success in solving questions that require contextual and ethical thinking [49]. This condition underscores the importance of viewing student performance as a dynamic process involving interrelated but cognitively distinct domains of knowledge.

4. Performance Distribution and Domain Interaction

However, not all students exhibit uneven trajectory patterns. Some students managed to maintain stable scores across all questions, reflecting consistency in understanding and the potential application of adaptive cognitive strategies. Additionally, some students showed a sudden spike in scores on Socio 2, which is suspected to be related to the relevance or familiarity of the question context. This supports the view that the connection between context and personal experience can enhance the quality of reasoning [50]. These findings reinforce the diagnostic value of spaghetti line plots, which present general trends and reveal individual dynamics often hidden in aggregate analyses. The implication is that assessment design needs to consider the order of questions to avoid detrimental sequential effects and integrate performance trajectory visualization approaches to provide more personalized, contextual, and process-focused formative feedback [51], [52].

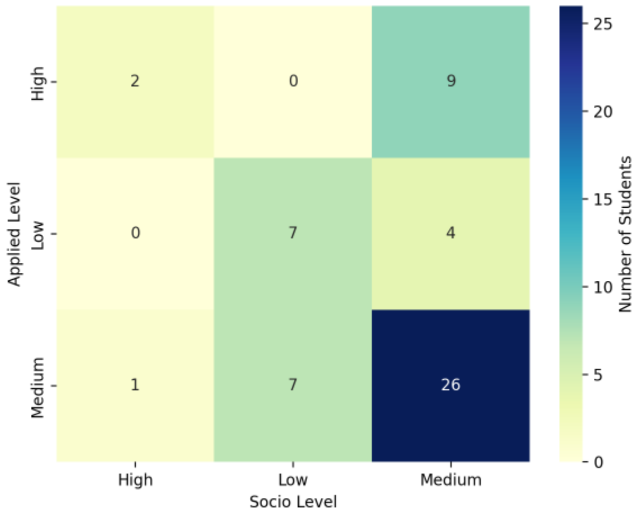

Figure 6. Heatmap of Applied and Socio Reasoning Levels.

Visualizing the interaction between student performance categories in Applied Reasoning and Socio-Chemical Reasoning provides significant insights into the distribution of abilities and transition patterns, as shown in Figure 6 (Heatmap Applied × Socio) and Figure 7 (Alluvial Diagram). Figure 6 shows that the highest density is in the Medium–Medium combination, where Applied scores range from 5–6 and Socio scores from 4–5. This finding indicates that most students are at a moderate performance level in both domains. However, there is also a fairly high concentration in the Low–Low region (Applied ≤4, Socio ≤3), indicating the presence of a group of students with overall low competencies who require special pedagogical attention [53]. This distribution creates two dominant poles—a group of students with relatively stable performance in the middle and a group needing deep intervention—which illustrates a sharp segmentation of abilities within the class.

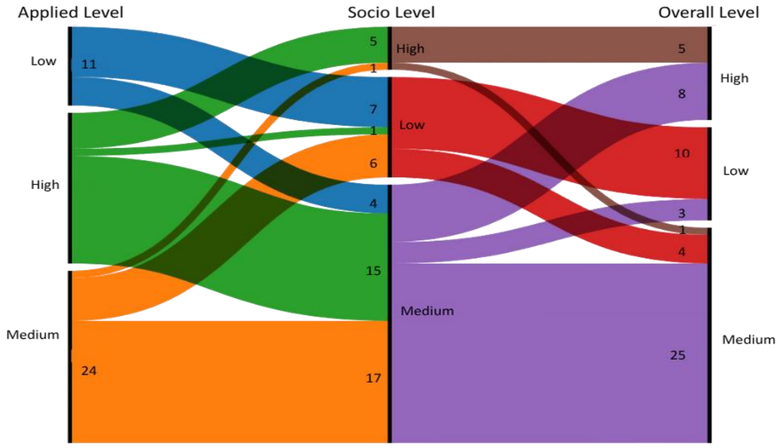

Figure 7. Alluvial Diagram of Reasoning Level Transitions.

Meanwhile, Figure 7 (Alluvial Diagram) shows the dynamics of performance trajectories between far more complex and non-linear levels. This diagram shows that some students with high scores in Applied Reasoning actually move to the Socio category at the moderate or low level, confirming that performance is not always synchronized across domains. Conversely, some students move from low to medium or high levels, and vice versa, reflecting the diversity of learning trajectories that conventional statistics cannot capture [54]. The Alluvial Diagram helps to visually reveal these transitions, showing that interactions between levels form a dynamic cognitive profile of students. Therefore, the evaluative implications are clear: adaptive and continuous cross-domain formative assessments must complement conventional summative assessments. With this approach, educators can more accurately understand students' performance development, design appropriate interventions, and ultimately support learning more responsive to students' cognitive variations [53], [54].

CONCLUSION

This study demonstrated that students' reasoning in chemistry is multidimensional, with significant intra- and inter-individual variations. Integrating Applied Reasoning and Socio-Chemical Reasoning in assessment highlighted a pattern of imbalanced proficiency across domains. Learning analytics tools like quadrant profiling, performance trajectory tracking, and cluster-based segmentation enable deeper insights into students' reasoning patterns and transitions. Theoretically, these findings extend existing models of chemical reasoning by emphasizing its non-linear and context-sensitive nature. Practically, the study implies the importance of adopting data-informed, integrated assessments to support instructional decision-making and curriculum development. However, the study was limited by its reliance on digital response formats, which may not capture all aspects of reasoning in diverse learner populations. Future research is recommended to validate these findings across different educational settings and explore the integration of socio-emotional reasoning indicators. This study contributes to developing responsive and inclusive assessment practices in science education.

REFERENCES

[1] D. King, "New perspectives on context-based chemistry education: Using a dialectical sociocultural approach to view teaching and learning," *Stud. Sci. Educ.*, vol. 48, no. 1, pp. 51–87, Mar. 2012, doi: 10.1080/03057267.2012.655037.

[2] D. G. Ramadhani, S. Yamtinah, S. Saputro, and S. Widoretno, "Analysis of the relationship between students' argumentation and chemical representational ability: A case study of hybrid learning oriented in the environmental chemistry course," *Chem. Teach. Int.*, vol. 5, no. 4, pp. 397–411, 2023, doi: 10.1515/cti-2023-0047.

[3] J. Pernaa, V. Kämppi, and M. Aksela, "Supporting the relevance of chemistry education through sustainable ionic liquids context: A research-based design approach," *Sustainability*, vol. 14, no. 10, p. 6220, 2022, doi: 10.3390/su14106220.

[4] J. L. Spencer et al., "Cultural relevance in chemistry education: Snow chemistry and the Iñupiaq community," *J. Chem. Educ.*, vol. 99, no. 1, pp. 363–372, 2021, doi: 10.1021/acs.jchemed.1c00480.

[5] J. E. Wissinger et al., "Integrating sustainability into learning in chemistry," *J. Chem. Educ.*, vol. 98, no. 4, pp. 1061–1063, 2021, doi: 10.1021/acs.jchemed.1c00284.

[6] V. Talanquer, "Importance of understanding fundamental chemical mechanisms," *J. Chem. Educ.*, vol. 95, no. 11, pp. 1905–1911, 2018, doi: 10.1021/acs.jchemed.8b00508.

[7] J. T. Laverty et al., "Characterizing college science assessments: The three-dimensional learning assessment protocol," *PLoS One*, vol. 11, no. 9, p. e0162333, 2016, doi: 10.1371/journal.pone.0162333.

[8] J. Sjöström, I. Eilks, and V. Talanquer, "Didaktik models in chemistry education," *J. Chem. Educ.*, vol. 97, no. 4, pp. 910–915, 2020, doi: 10.1021/acs.jchemed.9b01034.

[9] V. Talanquer, "Chemistry education: Ten facets to shape us," *J. Chem. Educ.*, vol. 90, no. 7, pp. 832–838, 2013, doi: 10.1021/ed300881v.

[10] J. H. Carmel, D. G. Herrington, L. A. Posey, J. S. Ward, A. M. Pollock, and M. M. Cooper, "Helping students to 'do science': Characterizing scientific practices in general chemistry laboratory curricula," *J. Chem. Educ.*, vol. 96, no. 3, pp. 423–434, 2019, doi: 10.1021/acs.jchemed.8b00912.

[11] C. Harris, J. Krajcik, J. W. Pellegrino, and A. H. DeBarger, "Designing knowledge-in-use assessments to promote deeper learning," *Educ. Meas. Issues Pract.*, vol. 38, no. 2, pp. 53–67, 2019, doi: 10.1111/emip.12253.

[12] J. R. Maeyer and V. Talanquer, "The role of intuitive heuristics in students' thinking: Ranking chemical substances," *Sci. Educ.*, vol. 94, no. 6, pp. 963–984, 2010, doi: 1 0.1002/sce.20397.

[13] L. Smith, J. R. Paddock, J. Vaughan, and D. W. Parkin, "Promoting nursing students' chemistry success in a collegiate active learning environment: 'If I have hope, I will try harder,'" *J. Chem. Educ.*, vol. 95, no. 11, pp. 1929–1938, 2018, doi: 10.1021/acs.jchemed.8b00201.

[14] Y. Jin, C. A. Rodriguez, L. Shah, and G. T. Rushton, "Examining the psychometric properties of the Redox Concept Inventory: A Rasch approach," *J. Chem. Educ.*, vol. 97, no. 12, pp. 4235–4244, 2020, doi: 10.1021/acs.jchemed.0c00479.

[15] K. Adu-Gyamfi and B. Anim-Eduful, "Interaction effect of gender, across school-type on upper-secondary students' development of experimental reasoning on organic qualitative analysis," *J. Balt. Sci. Educ.*, vol. 21, no. 3, pp. 351–364, 2022, doi: 10.33225/jbse/22.21.351.

[16] K. Adu-Gyamfi and I. A. Asaki, "Teachers' conceptual difficulties in teaching senior high school organic chemistry," *Contemp. Math. Sci. Educ.*, vol. 3, no. 2, p. ep22019, 2022, doi: 10.30935/conmaths/12382.

[17] L. N. Amsad, S. Liliasari, A. Kadarohman, and R. E. Sardjono, "Students' difficulties in solving synthesis organic compound problems," *Unnes Sci. Educ. J.*, vol. 8, no. 3, 2019, doi: 10.15294/usej.v8i3.32424.

[18] D. L. Danipog and M. B. Ferido, "Using art-based chemistry activities to improve students' conceptual understanding in chemistry," *J. Chem. Educ.*, vol. 88, no. 12, pp. 1610–1615, 2011, doi: 10.1021/ed100009a.

[19] B. Wei, "Reconstructing a school chemistry curriculum in the era of core competencies: A case from China," *J. Chem. Educ.*, vol. 96, no. 7, pp. 1359–1366, 2019, doi: 10.1021/acs.jchemed.9b00211.

[20] J. Fischmann, P. Verboon, and R. v. Geel, "Elementary tutorial for multilevel analysis using visualizations," 2021, doi: 10.31234/osf.io/7nzx2.

[21] L.-T. Tsai and C.-C. Yang, "Hierarchical effects of school-, classroom-, and student-level factors on the science performance of eighth-grade Taiwanese students," *Int. J. Sci. Educ.*, vol. 37, no. 8, pp. 1166–1181, 2015, doi: 10.1080/09500693.2015.1022625.

[22] J. Wu, Y.-H. Lee, and J. J. H. Lin, "Using iMCFA to perform the CFA, multilevel CFA, and maximum model for analyzing complex survey data," *Front. Psychol.*, vol. 9, 2018, doi: 10.3389/fpsyg.2018.00251.

[23] J. W. K. Yeung, "The dynamic relationships between educational expectations and science learning performance among students in secondary school and their later completion of a STEM degree," *Behav. Sci.*, vol. 14, no. 6, p. 506, 2024, doi: 10.3390/bs14060506.

[24] B. T. Moges, M. M. Gebremeskel, S. A. Tilwani, and Y. Assefa, "Student engagement in a differentiated higher education system in Ethiopia: A multilevel analysis," *J. Appl. Res. High. Educ.*, vol. 16, no. 5, pp. 1341–1354, 2024, doi: 10.1108/jarhe-11-2023-0507.

[25] O. Ersan and M. C. Rodriguez, "Socioeconomic status and beyond: A multilevel analysis of TIMSS mathematics achievement given student and school context in Turkey," *Large-Scale Assessments Educ.*, vol. 8, no. 1, 2020, doi: 10.1186/s40536-020-00093-y.

[26] B. Setiaji, P. H. Santoso, K. N. Aziz, H. Retnawati, and M. Khairudin, "Using multilevel modelling to evaluate science literacy and technology course of the Indonesian non-science students," *J. Pendidik. Ipa Indones.*, vol. 12, no. 1, pp. 96–111, 2023, doi: 10.15294/jpii.v12i1.41457.

[27] S. A. Priyambada, M. ER, B. N. Yahya, and T. Usagawa, "Profile-based cluster evolution analysis: Identification of migration patterns for understanding student learning behavior," *IEEE Access*, vol. 9, pp. 101718–101728, 2021, doi: 10.1109/access.2021.3095958.

[28] V. Revathi, K. B. Ghutugade, R. Vannapuram, and B. P. S. Prasanna, "K-means algorithm for clustering of

learners performance levels using machine learning techniques," *Rev. D Intell. Artif.*, vol. 35, no. 1, pp. 99–104, 2021, doi: 10.18280/ria.350112.

[29] C. Romero and S. Ventura, "Educational data mining and learning analytics: An updated survey," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 10, no. 3, 2020, doi: 10.1002/widm.1355.

[30] T. Yıldırım, "Trends in PhD theses in Turkish chemistry education (1999–2019)," *Eurasian J. Educ. Res.*, vol. 20, no. 89, pp. 1–40, 2020, doi: 10.14689/ejer.2020.89.10.

[31] M. Schultz, J. Lai, J. P. Ferguson, and S. Delaney, "Topics amenable to a systems thinking approach: Secondary and tertiary perspectives," *J. Chem. Educ.*, vol. 98, no. 10, pp. 3100–3109, 2021, doi: 10.1021/acs.jchemed.1c00203.

[32] S. York, R. Lavi, Y. J. Dori, and M. Orgill, "Applications of systems thinking in STEM education," *J. Chem. Educ.*, vol. 96, no. 12, pp. 2742–2751, 2019, doi: 10.1021/acs.jchemed.9b00261.

[33] B. Flynn et al., "Future directions for systems thinking in chemistry education: Putting the pieces together," *J. Chem. Educ.*, vol. 96, no. 12, pp. 3000–3005, 2019, doi: 10.1021/acs.jchemed.9b00637.

[34] P. G. Mahaffy et al., "Beyond 'inert' ideas to teaching general chemistry from rich contexts: Visualizing the chemistry of climate change (VC3)," *J. Chem. Educ.*, vol. 94, no. 8, pp. 1027–1035, 2017, doi: 10.1021/acs.jchemed.6b01009.

[35] D. J. C. Constable, C. Jiménez-González, and S. A. Matlin, "Navigating complexity using systems thinking in chemistry, with implications for chemistry education," *J. Chem. Educ.*, vol. 96, no. 12, pp. 2689–2699, 2019, doi: 10.1021/acs.jchemed.9b00368.

[36] H. Tümay, "Systems thinking in chemistry and chemical education: A framework for meaningful conceptual learning and competence in chemistry," *J. Chem. Educ.*, vol. 100, no. 10, pp. 3925–3933, 2023, doi: 10.1021/acs.jchemed.3c00474.

[37] M. A. Al-Hagery, M. A. Alzaid, T. S. Alharbi, and M. A. Alhanaya, "Data mining methods for detecting the most significant factors affecting students' performance," *Int. J. Inf. Technol. Comput. Sci.*, vol. 12, no. 5, pp. 1–13, 2020, doi: 1 0.5815/ijitcs.2020.05.01

[38] M. Usher, A. Hershkovitz, and A. Forkosh-Baruch, "From data to actions: Instructors' decision making based on learners' data in online emergency remote teaching," *Br. J. Educ. Technol.*, vol. 52, no. 4, pp. 1338–1356, 2021, doi: 10.1111/bjet.13108.

[39] Y. Tsai, O. Poquet, D. Gašević, S. Dawson, and A. Pardo, "Complexity leadership in learning analytics: Drivers, challenges and opportunities," *Br. J. Educ. Technol.*, vol. 50, no. 6, pp. 2839–2854, 2019, doi: 10.1111/bjet.12846.

[40] M. v. Geel, T. Keuning, A. J. Visscher, and J. Fox, "Assessing the effects of a school-wide data-based decision-making intervention on student achievement growth in primary schools," *Am. Educ. Res. J.*, vol. 53, no. 2, pp. 360–394, 2016, doi: 10.3102/0002831216637346.

[41] M. Botvin, A. Hershkovitz, and A. Forkosh-Baruch, "Data-driven decision-making in emergency remote teaching," *Educ. Inf. Technol.*, vol. 28, no. 1, pp. 489–506, 2022, doi: 10.1007/s10639-022-11176-4.

[42] A. S. Musyarofah, A. L. Anggraini, S. Sudarti, R. D. Handayani, M. Jamhari, and H. Haeruddin, "Analysis of the comparison of science literacy skills of students at MTS Nurul Huda Situbondo in solving PISA science problems," *Edukasia*, vol. 4, no. 2, pp. 2507–2516, 2023, doi: 10.62775/edukasia.v4i2.619.

[43] K. A. Kırkıç and F. Uludağ, "STEM attitudes of students as predictor of secondary school technology and design course achievement," *Probl. Educ. 21st Century*, vol. 79, no. 4, pp. 585–596, 2021, doi: 10.33225/pec/21.79.585.

[44] M. Mejía-Rodríguez and L. Kyriakidēs, "What matters for student learning outcomes? A systematic review of studies exploring system-level factors of educational effectiveness," *Rev. Educ.*, vol. 10, no. 3, 2022, doi: 10.1002/rev3.3374.

[45] A. Dewi and Y. Yahdi, "Research trends on socio-scientific issues in chemistry learning: A systematic review," *J. Pendidik. Mipa*, vol. 26, no. 1, pp. 457–475, 2025, doi: 10.23960/jpmipa.v26i1.pp457-475.

[46] N. Kravchenko and N. Mikheeva, "Theoretical construct of young learners' intercultural competence and its practical application," *Bull. South Ural State Univ. Ser. Educ. Educ. Sci.*, vol. 15, no. 3, pp. 32–41, 2023, doi: 10.14529/ped230303.

[47] F. Daniels, L. P. Fakude, N. S. Linda, and R. R. M. Modeste, "Nurse educators' experiences of case-based education in a South African nursing programme," *Curationis*, vol. 38, no. 2, 2015, doi: 10.4102/curationis.v38i2.1523.

[48] T. M. Moore et al., "Development of an abbreviated form of the Penn Line Orientation Test using large samples and computerized adaptive test simulation," *Psychol. Assess.*, vol. 27, no. 3, pp. 955–964, 2015, doi: 10.1037/pas0000102.

[49] S. Xu, A. Song, X. Li, J. Nie, C. Zhang, and Y. Wang, "Performance of the automated COBAS AmpliPrep/COBAS TaqMan HIV-1 test on a genetically diverse panel of specimens from China: Comparison to the COBAS Amplicor HIV-1 monitor test, v1.5," *Intervirology*, vol. 53, no. 4, pp. 221–228, 2010, doi: 10.1159/000299064.

[50] J. Afonso, R. Ramírez-Campillo, F. M. Clemente, F. Büttner, and R. Andrade, "The perils of misinterpreting and misusing 'publication bias' in meta-analyses: An education review on funnel plot-based methods," *Sport. Med.*, vol. 54, no. 2, pp. 257–269, 2023, doi: 10.1007/s40279-023-01927-9.

[51] L. Padilla, I. T. Ruginski, and S. H. Creem-Regehr, "Effects of ensemble and summary displays on interpretations of geospatial uncertainty data," *Cogn. Res. Princ. Implic.*, vol. 2, no. 1, 2017, doi: 10.1186/s41235-017-0076-1.

[52] R. P. Hickson, I. E. Annis, L. A. Killeya-Jones, and G. Fang, "Opening the black box of the group-based trajectory modeling process to analyze medication adherence patterns: An example using real-world statin adherence data," *Pharmacoepidemiol. Drug Saf.*, vol. 29, no. 3, pp. 357–362, 2019, doi: 10.1002/pds.4917.

[53] M. Stieff, M. Hegarty, and G. Deslongchamps, "Identifying representational competence with multi-representational displays," *Cogn. Instr.*, vol. 29, no. 1, pp. 123–145, 2011, doi: 10.1080/07370008.2010.507318.

[54] H. Wu and P. Shah, "Exploring visuospatial thinking in chemistry learning," *Sci. Educ.*, vol. 88, no. 3, pp. 465–492, 2004, doi: 10.1002/sce.10126.